

Lecture 1

This lecture provides an introduction to:

- **sets of numbers**
- **the "floor" and "ceiling" functions**
- **divisibility among the integers**

You should master the material contained in this lecture before moving on to the next lecture.

Part 1: Sets of Numbers

Sets and Numbers

A *set* is a collection of objects, which are called the *elements* of the set. We write

$$a \in X$$

to denote that a is an element of the set X. We also say that "a belongs to X", "a is in X", "a is a member of X" and "X includes a (as one of its elements)".

A set is often described by means of a pair of braces **{** and **}** enclosing the elements of the set, with the elements being separated by commas.

Example:
The set $X = \{a, b, c\}$ has elements a, b and c, so that we can write $a \in X$, $b \in X$ and $c \in X$.

End of example

If X and Y are sets, then we say that X is a *subset* of Y and write

$$X \subseteq Y$$

if every element of X is also in Y. We also say "X is contained in Y" or "Y contains X".

For example, if $\boldsymbol{X = \{2, 3, 4\}}$ and $\boldsymbol{Y = \{1, 2, 3, 4, 5\}}$ then $\boldsymbol{X \subseteq Y}$.

The sets of most interest to us have **numbers** as their elements. We describe the most important such sets below.

The set $\mathbb{N}$ consists of all *natural numbers* (or "counting numbers"), so that

$$\mathbb{N} = \{1, 2, 3, ...\}.$$

Note that this is our first example of an **infinite** set.

The set $\mathbb{Z}$ consists of all *integers* (or "whole numbers"), so that

$$\mathbb{Z} = \{..., -3, -2, -1, 0, 1, 2, 3, ...\}.$$

Note that $\mathbb{N} \subseteq \mathbb{Z}$. Since the positive numbers in $\mathbb{Z}$ are precisely the elements of $\mathbb{N}$, the natural numbers are also called the *positive integers*.

Certain **finite** sets of positive integers are of great importance in discrete mathematics. In particular, if $\boldsymbol{n \in \mathbb{N}}$ and $\boldsymbol{n \geq 2}$ then we define

$$\mathbb{Z}_n = \{0, 1, 2, 3, ..., n - 1\}$$

to be the set consisting of the first $\boldsymbol{n}$ non-negative integers. For example, $\mathbb{Z}_5 = \{0, 1, 2, 3, 4\}$. Note that, always, $\mathbb{Z}_n \subseteq \mathbb{Z}$.

When $\boldsymbol{n = 2}$, the set $\mathbb{Z}_n$ becomes $\mathbb{Z}_2 = \{0, 1\}$. This set is of crucial importance in Coding Theory. However, we will soon be discussing Boolean Algebra, where the set $\{0, 1\}$ again plays a key role. In the

Boolean Algebra context, this two-element set is denoted by $\mathbb{B}$, so that we write $\mathbb{B} = \{0, 1\}$. We sometimes call this the *binary* set, because it consists of the two binary digits 0 and 1.

If p and q are integers, and $q \neq 0$, then dividing p by q gives the expression

$$\frac{p}{q}$$

which is sometimes called the "quotient" of the two integers. (But beware: as we shall see, the same word is used in another context to mean the **integer part** of p/q.) Such an expression is called a *rational* number, as is any number which can be expressed in the form p/q. The set of all rational numbers is denoted by $\mathbb{Q}$.

Since every integer n can be expressed as $n/1$, all integers are rational numbers. So $\mathbb{Z} \subseteq \mathbb{Q}$, which when combined with an earlier result tells us that $\mathbb{N} \subseteq \mathbb{Z} \subseteq \mathbb{Q}$.

Some numbers are "irrational". For example, the number π ("pi") has no exact representation as a fraction (though it equals 22/7 to 2 decimal places, and 355/113 to 6 decimal places). Similarly, numbers like $\sqrt{2}$ (the positive square root of 2) cannot be expressed as fractions.

When we gather together all the rational numbers and all the irrational numbers, we have the set of numbers known as the *real numbers*. So, loosely speaking, vir-

tually everything is a real number. The set of all real numbers is denoted by $\mathbb{R}$. From the above discussion it is clear that $\mathbb{N} \subseteq \mathbb{Z} \subseteq \mathbb{Q} \subseteq \mathbb{R}$.

When we combine real numbers, using the familiar operations like addition, multiplication, subtraction and division, we almost always get real numbers as a result. But there are some things we aren't allowed to do. We can't divide by 0, and we can't take the square root of a negative number.

Worked Example:
We decide which sets have the following numbers among their elements: $\mathbf{0}$, $\mathbf{-3}$, $\mathbf{\sqrt{8}}$, $\mathbf{\sqrt{9}}$, $\mathbf{-5/2}$, $\mathbf{\sqrt{-9}}$

Since 0 is an integer, but not a positive integer, it is in $\mathbb{Z}$ and therefore $\mathbb{Q}$ and therefore $\mathbb{R}$. But it isn't in $\mathbb{N}$. Similarly, $\mathbf{-3}$ is in $\mathbb{Z}$, $\mathbb{Q}$ and $\mathbb{R}$ but not $\mathbb{N}$.

Since $\mathbf{8}$ is positive, its square root exists. So $\mathbf{\sqrt{8}}$ is in $\mathbb{R}$. But it is not in any of the other three sets.

Since $\mathbf{\sqrt{9} = 3}$, which is a positive integer, $\mathbf{\sqrt{9}}$ is in all four of the sets.

Since $\mathbf{-5/2}$ is rational but is not an integer, it is only in $\mathbb{Q}$ and $\mathbb{R}$.

Since $\mathbf{-9}$ is negative, it has no real square roots. So $\mathbf{\sqrt{-9}}$ is not in $\mathbb{R}$, and hence is not in any of the other three sets either.

End of worked example

Formative Assessment

You should now do as many practice exercises as necessary to establish that you can correctly determine which of the sets $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$ and $\mathbb{R}$ have any given number as one of their elements.

End of formative assessment

Floors and Ceilings

There are two important functions that relate real numbers to integers. These are the "floor" and the "ceiling" functions, which we now define.

If $\boldsymbol{x} \in \mathbb{R}$, then the *floor* $\lfloor \boldsymbol{x} \rfloor$ of $\boldsymbol{x}$ is the greatest integer less than or equal to $\boldsymbol{x}$. When $\boldsymbol{x}$ is positive, this is also called the *integer part* of $\boldsymbol{x}$.

Example:
$\lfloor \mathbf{1.4} \rfloor = \mathbf{1}$, $\lfloor \boldsymbol{\pi} \rfloor = \mathbf{3}$, $\lfloor \sqrt{\mathbf{3}} \rfloor = \mathbf{1}$, $\lfloor \mathbf{4} \rfloor = \mathbf{4}$, $\lfloor \mathbf{-5} \rfloor = \mathbf{-5}$, $\lfloor \mathbf{-1.4} \rfloor = \mathbf{-2}$, $\lfloor \mathbf{-\pi} \rfloor = \mathbf{-4}$, $\lfloor \mathbf{-\sqrt{3}} \rfloor = \mathbf{-2}$.

End of example

Similarly, the *ceiling* $\lceil \boldsymbol{x} \rceil$ of $\boldsymbol{x}$ is the smallest integer greater than or equal to $\boldsymbol{x}$. When $\boldsymbol{x}$ is negative, this is the "integer part" of $\boldsymbol{x}$.

Example:
$\lceil \mathbf{1.4} \rceil = \mathbf{2}$, $\lceil \boldsymbol{\pi} \rceil = \mathbf{4}$, $\lceil \sqrt{\mathbf{3}} \rceil = \mathbf{2}$, $\lceil \mathbf{4} \rceil = \mathbf{4}$, $\lceil -\mathbf{5} \rceil = -\mathbf{5}$, $\lceil -\mathbf{1.4} \rceil = -\mathbf{1}$, $\lceil -\boldsymbol{\pi} \rceil = -\mathbf{3}$, $\lceil -\sqrt{\mathbf{3}} \rceil = -\mathbf{1}$.

End of example

Note that if $\boldsymbol{x}$ is an integer then $\lceil \boldsymbol{x} \rceil = \lfloor \boldsymbol{x} \rfloor$ $(= \boldsymbol{x})$, but if $\boldsymbol{x}$ is not an integer then $\lceil \boldsymbol{x} \rceil = \lfloor \boldsymbol{x} \rfloor + \mathbf{1}$.

Exercise:
Find the floor and ceiling of $-\mathbf{1/2}$.

End of exercise

Formative Assessment

You should now do as many practice exercises as necessary to establish that you can correctly find the floor and ceiling of any real number.

End of formative assessment

Divisibility

The integers form a closed system under $+$ and $\cdot$, in the sense that the sum or product of two integers is again an integer.

It is natural then to ask about the two operations that are the inverses of $+$ and $\cdot$, namely $-$ and $\div$ respectively.

Subtraction causes few difficulties, since it still produces integers. However, subtraction can take us from the set $\mathbb{N}$ into the set $\mathbb{Z}$.

Division is an entirely different matter. Dividing 3 by 6, for example, gives the rational number $\mathbf{1/2}$ which is not an integer. Furthermore, division by 0 never even produces a real number.

We need to develop notation and terminology to describe divisibility relations among the integers, and to deal with the various possible outcomes of division. The critical question is this. When does division produce an integral result, and what happens when it doesn't?

Let $\boldsymbol{m}$ and $\boldsymbol{n}$ be positive integers. Multiplying together these two numbers gives their **product** $\boldsymbol{mn}$.

If $\boldsymbol{m = n}$ then $\boldsymbol{mn = m^2}$. Continuing to multiply by $\boldsymbol{m}$ gives us the positive powers of $\boldsymbol{m}$: $\boldsymbol{m^1}$ $(= \boldsymbol{m})$, $\boldsymbol{m^2}$, $\boldsymbol{m^3}$, and so on. (There also exist the zero power $\boldsymbol{m^0}$, which is defined to be 1, and the negative powers $\boldsymbol{m^{-1} = 1/m}$, $\boldsymbol{m^{-2} = 1/m^2}$,

and so on.)

Put $\boldsymbol{t = mn}$. There are various ways to describe the relationship between $\boldsymbol{t}$ and each of $\boldsymbol{m}$ and $\boldsymbol{n}$. We will just discuss $\boldsymbol{t}$ and $\boldsymbol{m}$. (Similar remarks can be made about $\boldsymbol{t}$ and $\boldsymbol{n}$.)

The integer $\boldsymbol{m}$ is a *factor* or a *divisor* of $\boldsymbol{t = mn}$. We also say that $\boldsymbol{m}$ *divides* $\boldsymbol{t}$ and write

$$\boldsymbol{m|t.} \tag{1}$$

Here, the vertical bar (|) represents the verb "divides". Rewriting (1) gives

$$\boldsymbol{m|(mn).} \tag{2}$$

The statement of (1) also means that $\boldsymbol{t}$ *is divisible by* $\boldsymbol{m}$. We also say that $\boldsymbol{t}$ is a *multiple* of $\boldsymbol{m}$.

It is apparent that

$$\boldsymbol{m|t} \text{ if } \frac{\boldsymbol{t}}{\boldsymbol{m}} \text{ is an integer.} \tag{3}$$

So $\boldsymbol{m|(mn)}$ because $\dfrac{\boldsymbol{mn}}{\boldsymbol{m}} = \boldsymbol{n}$.

A partial converse of (3) is also true:

if $\boldsymbol{m|t}$ and $\boldsymbol{m \neq 0}$ then $\dfrac{\boldsymbol{t}}{\boldsymbol{m}}$ is an integer.

But note that

$$\mathbf{0|0} \tag{4}$$

even though $\mathbf{0/0}$ is indeterminate. We know (4) is true, because $\mathbf{0 \cdot 3 = 0}$ (for example) so that both $\mathbf{0}$ and $\mathbf{3}$ are divisors of $\mathbf{0}$.

By the *factors* of $\boldsymbol{t}$ we mean all those positive integers which divide $\boldsymbol{t}$. The smallest

of these is 1, and the largest is $\boldsymbol{t}$. They are also called the *divisors* of t. (More precisely, they are the *positive integral* divisors of t.)

A *factorisation* of an integer $\boldsymbol{t}$ is a representation of $\boldsymbol{t}$ as the product of some of its factors. So if $\boldsymbol{t}$ is the product of two integers $\boldsymbol{m}$ and $\boldsymbol{n}$ then

$$\boldsymbol{t = mn} \tag{5}$$

is one factorisation. A special case is

$$\boldsymbol{t = 1 \cdot t} \tag{6}$$

which is known as the *trivial* factorisation of $\boldsymbol{t}$.

Worked Example:
The factors of 24 are 1, 2, 3, 4, 6, 8, 12 and 24. The easiest way to find them is to write down 1 as the smallest factor, and then divide 24 by 1 to give 24 as the largest factor. Then run through the other small integers 2, 3, 4 and so on, writing down any which divide 24. For each such small factor, dividing 24 by that factor gives a large factor of 24. For example, $24/2 = 12$. Continuing this process, the "small" factors eventually meet up with the "large" factors, so that there are no more factors to be found. Using this strategy, the order in which we would discover the factors is as

follows: 1, 24, 2, 12, 3, 8, 4, 6. Trying out 5, we find it is not a factor. The next one to try would be 6, but it has already been recorded as a factor. So we are finished.

End of worked example

Formative Assessment

You should now do as many practice exercises as necessary to establish that you can correctly find all the integral factors of a given positive integer.

End of formative assessment

"Prime" numbers can only be trivially factorised. We will study them in the next lecture.

Before proceeding to the next lecture, it would be instructive for you to spend some time trying to:

- identify some 2-digit numbers that don't seem to have many factors;
- revise your strategies for adding two fractions with different denominators.

Lecture 2

This lecture provides an introduction to:

- **prime numbers and the prime power factorisation**
- **quotients and remainders**
- **greatest common divisors and least common multiples**

You should master the material contained in this lecture before moving on to the next lecture.

Prime Numbers

A *prime number* is a positive integer which can only be factorised in one way and whose only factorisation represents it as the product of two distinct positive integers (namely itself and 1). Equivalently, it is an integer greater than 1 whose only factorisation is the trivial one.

Example:
The first ten prime numbers are

2, 3, 5, 7, 11, 13, 17, 19, 23 and 29.

Note that 2 is the only **even** prime number.

End of example

Exercise:
Write down the prime numbers between 30 and 40.

End of exercise

A positive integer $> \mathbf{1}$ which is not prime is said to be *composite*. Note that 1 is usually regarded as being neither prime nor composite.

Example:
The first ten composite numbers are

4, 6, 8, 9, 10, 12, 14, 15, 16 and 18.

Note that every even number greater than 2 is composite.

End of example

Prime numbers play an important role in the factorisation of integers. This is why they have been used in recent times in the development of cryptosystems for information security.

To test a number t to see if it's prime, we only need to check all the divisors less than or equal to its square root. In fact, after a little thought it is clear that we only need to check the prime numbers less than or equal to the square root. If none of them divides t, then t is a prime number.

A key relationship between prime numbers and other positive integers is given in the next theorem.

Theorem:
Every natural number has a unique prime power factorisation.

End of theorem

Our previous discussion leads fairly naturally to this result. To display a natural number t in this way, we identify its prime factors. For each prime factor p, we find the highest power p^s of p which also divides t. Then the product of all such powers is equal to t.

Worked Example:
Consider the integer $\boldsymbol{t = 90}$. We want to express it as a product of powers of prime numbers. Firstly, we use interpolation to estimate $\sqrt{\mathbf{90}}$. Locating the closest square numbers on either side of 90, we find that they are 81 and 100.

Now since

$$\mathbf{81 < 90 < 100}$$

we can conclude that

$$\sqrt{\mathbf{81}} < \sqrt{\mathbf{90}} < \sqrt{\mathbf{100}};$$

that is,

$$\mathbf{9 < \sqrt{90} < 10.}$$

This is because the square root function is an *increasing* function: as $\boldsymbol{x}$ increases, so does $\sqrt{\boldsymbol{x}}$.

Since 90 is roughly equidistant from 81 and 100, $\sqrt{\mathbf{90}}$ is *very* roughly equidistant from 9 and 10. So $\sqrt{\mathbf{90}} \approx \mathbf{9\frac{1}{2}}$.

This means that the prime numbers less than or equal to $\sqrt{\mathbf{90}}$ are 2, 3, 5 and 7. By inspection, we find that the first three are factors of 90.

Since 90 has some prime factors, it is not itself a prime number. (Actually it's easy to see this from the fact that 90 is even. But sometimes it takes a while to determine whether or not a number is prime.)

Now we list all the positive integral divisors of 90 (using the method discussed towards the end of Lecture 1). They are 1, 2, 3, 5, 6, 9, 10, 15, 18, 30, 45 and 90. Looking through this list, we see that in this case there are no other prime numbers except the three we already identified. That is, they are the only prime divisors of 90. (Note that sometimes a number $\boldsymbol{n}$ could have a prime divisor greater than $\sqrt{\boldsymbol{n}}$. But always the prime divisors are less than or equal to $\boldsymbol{n/2}$.)

We have identified 2, 3 and 5 as the prime divisors of 90. So we examine their powers.

Powers of **2**: $\mathbf{2|90}$, but $\mathbf{2^2}$ doesn't.
Powers of **3**: $\mathbf{3|90}$, $\mathbf{3^2|90}$, but $\mathbf{3^3}$ doesn't.
Powers of **5**: $\mathbf{5|90}$, but $\mathbf{5^2}$ doesn't.

Now that we have found the largest prime power divisors of **90** to be **2**, $\mathbf{3^2}$ and **5**, we can immediately write down that **90** is the product of these powers:

$$\mathbf{90 = 2 \cdot 3^2 \cdot 5.}$$

All that remains to be done is to check that the product on the right-hand side of the equation does indeed equal **90** (just in case we've made a mistake somewhere). It turns out that the equality is correct.

End of worked example

Formative Assessment

You should now do as many practice exercises as necessary to establish that you can correctly find the prime power factorisation of a given integer.

End of formative assessment

Quotients and Remainders

We now proceed to consider what happens if we divide one integer $\boldsymbol{a}$ by another integer $\boldsymbol{b}$ when $\boldsymbol{b}$ is not a factor of $\boldsymbol{a}$.

Let $\boldsymbol{c}$ be the largest integer below $\boldsymbol{a}$ which is a multiple of $\boldsymbol{b}$. So

$$\boldsymbol{c = qb}$$

for some $\boldsymbol{q} \in \mathbb{N}$. Put $\boldsymbol{r = a - c}$. So $\boldsymbol{0 \leq r < b}$. Then

$$\boldsymbol{a = c + r}$$

that is,

$$\boldsymbol{a = qb + r.} \tag{1}$$

Here, $\boldsymbol{q}$ is called the *quotient*, and $\boldsymbol{r}$ the *remainder*, *after division* of $\boldsymbol{a}$ by $\boldsymbol{b}$.

This process is summarised in what is known as the *division algorithm.*

Division Algorithm:

Let $\boldsymbol{a}$ and $\boldsymbol{b}$ be integers, and assume that $\boldsymbol{b}$ is positive. Then there exist integers $\boldsymbol{q}$ and $\boldsymbol{r}$ such that

$$\boldsymbol{a = qb + r}$$

where $\mathbf{0 \leq r < b}$.

End of algorithm

Remark.

The word "quotient" is sometimes used to mean the real number $\boldsymbol{x}$ such that $\boldsymbol{xb = a}$, so that $\boldsymbol{x = a/b}$. In the present situation, the quotient is the *floor* $\boldsymbol{\lfloor a/b \rfloor}$ of $\boldsymbol{a/b}$.

Worked Example:
We use the Division Algorithm to obtain the quotient and the remainder after each of the following divisions: dividing 3 by 7, $\mathbf{-4}$ by 3 and 119 by 5.

$$\mathbf{3 = 0 \cdot 7 + 3} \quad \therefore \boldsymbol{q = 0}, \boldsymbol{r = 3}$$
$$\mathbf{-4 = -2 \cdot 3 + 2} \quad \therefore \boldsymbol{q = -2}, \boldsymbol{r = 2}$$
$$\mathbf{119 = 23 \cdot 5 + 4} \quad \therefore \boldsymbol{q = 23}, \boldsymbol{r = 4}$$

End of example

Exercise:
Find the quotient $\boldsymbol{q}$ and remainder $\boldsymbol{r}$ when 30 is divided by 4.

End of exercise

Common Factors and Common Multiples

We now consider two other important binary operations defined on the set $\mathbb{N}$ of all natural numbers. These are the *greatest common divisor* and the *least common multiple* operations.

The names are self-explanatory. Let a and b be positive integers. A **common** divisor is an integer which divides both a and b. Clearly, one such common divisor is the smallest positive integer, 1. But often there are larger integers which divide both a and b. The largest of these is the **greatest** common divisor, or *gcd*. It is also called the **highest common factor** of a and b.

Analogously, the least common multiple of a and b can be obtained in the following way. A **common** multiple of a and b is an integer which is a multiple of a and is also a multiple of b. Clearly, one such common multiple is their product ab. But often there are smaller integers which are multiples of both a and b. The smallest of these is the **least** common multiple, or *lcm*. It is the number which is used as the "lowest common denominator" when we add two or more fractions.

We now proceed to look at how to find the **gcd** and the **lcm** of two given numbers.

For relatively small values of a and b the

gcd can be obtained by simply listing the divisors of both numbers.

Worked Example:
We find the **gcd** of 18 and 24. The factors of 18 are 1, 2, 3, 6, 9 and 18. The factors of 24 are 1, 2, 3, 4, 6, 8, 12 and 24. So the common factors are 1, 2, 3 and 6, of which the largest is 6.

End of worked example

Exercise:
Find the **gcd** of 15 and 24.

End of exercise

The method illustrated above is not really satisfactory when $\boldsymbol{a}$ and $\boldsymbol{b}$ are large. Fortunately there is a method available for dealing with such situations. It is called the *Euclidean algorithm*.

Euclidean Algorithm:
Let $\boldsymbol{a}$ and $\boldsymbol{b}$ be natural numbers. Suppose we use the division algorithm repeatedly as follows:

$$
\begin{aligned}
a &= q_0 b + r_1 \\
b &= q_1 r_1 + r_2 \\
r_1 &= q_2 r_2 + r_3 \\
&\vdots \\
r_{n-1} &= q_n r_n + r_{n+1} \\
r_n &= q_{n+1} r_{n+1}
\end{aligned}
$$

We stop when the remainder becomes zero. Then the previous remainder — the last nonzero remainder, r_{n+1} — is the **gcd** of a and b.

End of algorithm

At first sight this algorithm may appear somewhat formidable, but in fact it is quite easy to use if you follow the procedure illustrated in the next worked example.

Worked Example:
We obtain the **gcd** for 94 and 12:

$$
\begin{aligned}
94 &= 7 \cdot 12 + 10 \\
12 &= 1 \cdot 10 + 2 \\
10 &= 6 \cdot 2
\end{aligned}
$$

So the **gcd** is 2.

To keep track of the remainders, we used a table like this.

94
12
10
2
0

The table displays the two integers (with the bigger one first), followed by the remainders in the order they are calculated.

End of worked example

The key idea is to keep a column of the remainders, but with the two original integers listed first in decreasing order. After calculating the first remainder, add it to the bottom of the column and cross out the number at the top. Then apply the Division Algorithm to the two numbers at the bottom of the column. Continue this process. Each time a new remainder is produced, cross out the remaining number highest in the column. Eventually the last remainder in the column is 0. The number just above it is the required greatest common divisor.

If we record and correctly process all the quotients and remainders produced, the Euclidean algorithm can be used to find the multiplicative inverse of an integer in a *finite field*. This is a crucial step in implementing certain methods of cryptanalysis. The details of this application can be found in a specialised course on cryptography.

When $\boldsymbol{a}$ and $\boldsymbol{b}$ are relatively small, we can

find the **lcm** by simply listing the multiples of each number, possibly right up to their product.

Worked Example:
We find the **lcm** of 8 and 10. The multiples of 8 are 8, 16, 24, 32, 40, ..., 80. The multiples of 10 are 10, 20, 30, 40, ..., 80. So the least common multiple is 40.

End of worked example

Exercise:
Find the **lcm** of 9 and 12.

End of exercise

A more systematic way of finding the **lcm** is provided by the strong relationship between the **lcm** and the **gcd**. It is summarised in the following theorem.

Theorem:
If $\boldsymbol{a}$ and $\boldsymbol{b}$ have **gcd** $\boldsymbol{d}$, then the **lcm** of $\boldsymbol{a}$ and $\boldsymbol{b}$ is

$$\boldsymbol{M = \frac{ab}{d}.}$$

That is, we divide the **product** by the **gcd** to obtain the **lcm**.

End of theorem

Remark. It is fairly easy to show, using the facts that $\boldsymbol{a}$ and $\boldsymbol{b}$ are multiples of $\boldsymbol{d}$, that $\boldsymbol{ab/d}$ is a common multiple of $\boldsymbol{a}$ and $\boldsymbol{b}$. It is a bit more difficult to show that it is al-

ways the **least** common multiple. We omit the details of the proof.

Note also that the prime power factorisations of $\boldsymbol{a}$ and $\boldsymbol{b}$ can be used to construct the **gcd** and the **lcm** of $\boldsymbol{a}$ and $\boldsymbol{b}$.

Example:
We apply this to an earlier (worked) example in which we found the **gcd** of 18 and 24 to be 6. So the **lcm** is

$$M = \frac{\mathbf{18 \cdot 24}}{\mathbf{6}}.$$

Cancelling out the 6 in the denominator with the 6 hiding inside 18 as one of its factors, we get

$$M = \frac{\mathbf{3 \cdot 24}}{\mathbf{1}}$$

so that the **lcm** is $\boldsymbol{M = 72}$.

End of example

Note that we can also find the **gcd** and **lcm** of several positive integers, rather than just two. Straightforward modifications of the methods discussed here apply to the more general case. We omit the details.

Formative Assessment

You should now do as many practice exercises as necessary to establish that you can correctly find the greatest common divisor and the least common multiple of two integers.

End of formative assessment

Lecture 3

This lecture provides an introduction to:

- **sets, cardinality and subsets**
- **Cartesian products**

You should master the material contained in this lecture before moving on to the next lecture.

Part 3: Sets and Functions

Subsets and Cardinality

The concept of a *set* is fundamental to discrete mathematics. We have seen that certain kinds of important numbers form sets. Often we can specify a set by describing some property that a typical element of the set has.

Example:
The set of all even numbers can be described as follows:

$$\{x \in \mathbb{Z} : x/2 \in \mathbb{Z}\}$$

We read this as *the set of all* x *in* $\mathbb{Z}$ *such that* $x/2$ *is in* $\mathbb{Z}$.

End of example

But sets occur more widely than just as sets of numbers.

Whenever we consider a collection of objects (or even people), we usually have a set.

We have seen that some sets such as $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$ and $\mathbb{R}$ are *infinite*. So we know something about how large a set can be. But how small can a set be?

Consider the two-element set $\{a, b\}$. Re-

moving the element $\boldsymbol{b}$ leaves a 1-element set $\boldsymbol{\{a\}}$. Note that no commas are needed to separate elements, if a set has only one element. A one-element set is also called a *singleton*.

Now we can take this process of removing elements one step further. Removing the element $\boldsymbol{a}$ leaves just the braces:

$$\boldsymbol{\{\}}$$

This notation describes a set with no elements. It is called the *null set* or the *empty set*.

An alternative notation for the null set is provided by the symbol

$$\boldsymbol{\varnothing}$$

which is similar to the letter Ø of the Norwegian alphabet. This should not be confused with the Greek letter *phi*, denoted by $\boldsymbol{\Phi}$ as a capital letter and by $\boldsymbol{\phi}$ in lower case. More importantly, it should not be confused with the number *zero* or *nought*, which in a computer science context often looks similar to this: $\boldsymbol{\emptyset}$. Note that the null-set symbol is rounder in shape, whereas the zero symbol is narrower and more oval or elliptical.

So there are two common ways to represent the null set:

$$\begin{aligned} \boldsymbol{\varnothing} &= \boldsymbol{\{\}} \\ &= \textbf{the null (empty) set} \end{aligned}$$

But don't make the mistake of putting the two representations together and regarding the combination as having the same meaning. That is, the set

$$\{\varnothing\}$$

certainly exists, but it is **not** the null set! In fact, it is a one-element set. This will be explained further when we discuss power sets.

We have briefly discussed the idea of a *subset* in the past. It needs to be remarked that *every set is a subset of itself*, and that *the null set is a subset of every set* (including itself).

Example:

If $\boldsymbol{X = \{a, b, c\}}$ then

$$\boldsymbol{\varnothing \subseteq \{a\} \subseteq \{a, b\} \subseteq \{a, b, c\} \subseteq X}$$

although we would normally replace the last containment symbol ($\subseteq$) by an *equals* sign ($=$).

End of example

In general, *containment* includes the possibility of *equality* as a special case. That's why the bottom part of the containment symbol has the appearance of an equality symbol. If we want to emphasise that one set $\boldsymbol{X}$ is contained in another set $\boldsymbol{Y}$ but that they are not equal, we say that $\boldsymbol{X}$ is *properly* contained in $\boldsymbol{Y}$ or is a *proper*

subset of $\boldsymbol{Y}$. This is shown as follows:

$$\boldsymbol{X} \subset \boldsymbol{Y} \text{ or } \boldsymbol{X} \subsetneq \boldsymbol{Y}.$$

The number of elements in a set, which can reasonably be called the "size" of the set, is more formally called the *cardinality* of the set. The cardinality of the set $\boldsymbol{X}$ is denoted by $|\boldsymbol{X}|$.

Example:
If $\boldsymbol{X}$ is the binary set $\mathbb{B} = \{0, 1\}$, then:

$$|\boldsymbol{X}| = |\mathbb{B}| = 2$$

End of example

Example:
The cardinality of any singleton is 1. For example:

$$|\{3\}| = 1$$

End of example

Example:
The null set has cardinality nought:

$$|\varnothing| = 0$$

End of example

It is fairly easy to see that

$$\text{if } \boldsymbol{X} \subseteq \boldsymbol{Y} \text{ then } |\boldsymbol{X}| \leq |\boldsymbol{Y}|.$$

Furthermore, **if $\boldsymbol{Y}$ is finite** we can say that

$$\text{if } \boldsymbol{X} \subset \boldsymbol{Y} \text{ then } |\boldsymbol{X}| < |\boldsymbol{Y}|.$$

However, this is **not** true for infinite sets, as we will explain later. This illustrates the fact that the notion of cardinality is more complicated for infinite sets. For example, we don't simply say that "the set $\mathbb{Z}$ of all integers has cardinality infinity". One difficulty lies in the fact that *two infinite sets need not have the same cardinality.* To put it another way, *not all infinite numbers are equal.* It turns out that there is not just one infinity, but rather there are various infinities of **different** infinite size. This will be explored further in a later lecture.

For a small finite set $\boldsymbol{X}$ it is easy to write down a complete list of all the subsets of $\boldsymbol{X}$. Suppose that $\boldsymbol{X}$ has $\boldsymbol{n}$ elements. We start by writing down the 0-element subset $\boldsymbol{\varnothing}$, then (if $\boldsymbol{n > 0}$) all the 1-element subsets, then (if $\boldsymbol{n > 1}$) all the 2-element subsets, and so on, finishing with the set $\boldsymbol{X}$ itself.

Worked Example:
Here are the subsets of $\boldsymbol{X = \{a, b, c\}}$:

0-element: $\varnothing$
1-element: $\{a\}$, $\{b\}$, $\{c\}$
2-element: $\{a, b\}$, $\{a, c\}$, $\{b, c\}$
3-element: $\{a, b, c\} = X$

End of worked example

Now we have already mentioned that a set can have any kind of objects as its members. In particular, those members could themselves be sets! In that situation we have a *set of sets*.

Example:
The set $\{\varnothing, \mathbb{N}, \mathbb{R}\}$ is a 3-element set, each of whose members is a set.

End of example

One very important way to create a set of sets is to gather together all of the subsets of some given set X. This creates **the set of all subsets of X**, which is known as the *power set* of X. It is denoted by $\mathcal{P}(X)$.

Example:
The 8 subsets of $X = \{a, b, c\}$ listed in the last worked example are the 8 elements of the power set $\mathcal{P}(X)$.

End of worked example

Observe that in the last worked example

we produced a power set of cardinality 8 from a set of cardinality 3. In this context, the key relationship between these two cardinalities is that $\mathbf{2^3 = 8}$.

More generally, if a set $\boldsymbol{X}$ has $\boldsymbol{n}$ elements then its power set has $\mathbf{2}^{\boldsymbol{n}}$ elements. (This will be proved in a later lecture.) So:

$$|\mathcal{P}(\boldsymbol{X})| = \mathbf{2}^{|\boldsymbol{X}|}$$

This is the reason for the name "power set". The cardinality of the power set is always a power of $\mathbf{2}$ (with the index of the power being the cardinality $|\boldsymbol{X}|$ of $\boldsymbol{X}$). In fact, to emphasise this relationship, the alternative notation

$$\mathbf{2}^{\boldsymbol{X}}$$

is sometimes used for the power set $\mathcal{P}(\boldsymbol{X})$. That is,

$$\mathcal{P}(\boldsymbol{X}) = \mathbf{2}^{\boldsymbol{X}}.$$

However, many people have difficulty interpreting the expression $\mathbf{2}^{\boldsymbol{X}}$ as a set rather than as a number. So usually we won't use that notation.

Worked Example:
Let $\boldsymbol{X} = \{\boldsymbol{a}, \boldsymbol{b}\}$. Then:

$$\mathcal{P}(\boldsymbol{X}) = \{\varnothing, \{\boldsymbol{a}\}, \{\boldsymbol{b}\}, \boldsymbol{X}\}$$

Note that

$$|\mathcal{P}(\boldsymbol{X})| = 4 = \mathbf{2}^{\mathbf{2}}.$$

End of worked example

Worked Example:
Let $X = \{a\}$. Then:

$$\mathcal{P}(X) = \{\varnothing, X\}$$

Note that

$$|\mathcal{P}(X)| = 2 = 2^1.$$

End of worked example

Worked Example:
Let $X = \varnothing$. Then:

$$\mathcal{P}(X) = \{\varnothing\}$$

Note that

$$|\mathcal{P}(X)| = 1 = 2^0.$$

End of worked example

The last worked example explains why the set $\{\varnothing\}$ is not the null set. Rather, it is the **power set** of the null set. It is a 1-element set, having the null set as its only element.

Put simply, if you open and close braces and write down **anything** between them (even the symbol for the null set), then you have put something in to the set, so that it is **not empty**!

Formative Assessment

You should now do as many practice exercises as necessary to establish that you can correctly determine the cardinality and the subsets of various sets.

End of formative assessment

Cartesian Products

Let $\boldsymbol{X}$ and $\boldsymbol{Y}$ be sets. Let $\boldsymbol{x} \in \boldsymbol{X}$ and $\boldsymbol{y} \in \boldsymbol{Y}$. Then the object $\boldsymbol{(x, y)}$ is an *ordered pair* from $\boldsymbol{X}$ and $\boldsymbol{Y}$ (in that order).

Example:
If $\boldsymbol{X = \{a, b, c\}}$ and $\boldsymbol{Y = \{0, 1\}}$ then the ordered pairs from $\boldsymbol{X}$ and $\boldsymbol{Y}$ are:

$$\boldsymbol{(a, 0), (a, 1), (b, 0), (b, 1), (c, 0), (c, 1)}$$

In contrast, the ordered pairs from $\boldsymbol{Y}$ and $\boldsymbol{X}$ (in that order) are:

$$\boldsymbol{(0, a), (0, b), (0, c), (1, a), (1, b), (1, c)}$$

End of example

Ordered pairs are often used to indicate the position of a point in two dimensions. For example, an ordered pair whose first entry is latitude and whose second is longitude represents the position of a point on the Earth's surface.

The set of all ordered pairs from $\boldsymbol{X}$ and $\boldsymbol{Y}$ is denoted by

$$\boldsymbol{X} \times \boldsymbol{Y}$$

and it is called the *Cartesian product* of $\boldsymbol{X}$ and $\boldsymbol{Y}$ (after the 17th-century French mathematician and philosopher René Descartes). So we can write:

$$\boldsymbol{X} \times \boldsymbol{Y} = \{(\boldsymbol{x}, \boldsymbol{y}) : \boldsymbol{x} \in \boldsymbol{X}, \boldsymbol{y} \in \boldsymbol{Y}\}$$

Note that there is no multiplication taking place here. The sets $\boldsymbol{X}$ and $\boldsymbol{Y}$ are not being "multiplied" together. So the term "Cartesian product", and the notation, may be misleading. The reason they are used is as follows: *the number of ordered pairs* from $\boldsymbol{X}$ and $\boldsymbol{Y}$ equals the product of the cardinalities of $\boldsymbol{X}$ and $\boldsymbol{Y}$. So

$$|\boldsymbol{X} \times \boldsymbol{Y}| = |\boldsymbol{X}| \times |\boldsymbol{Y}|$$

although we would normally write this statement as

$$|\boldsymbol{X} \times \boldsymbol{Y}| = |\boldsymbol{X}| \cdot |\boldsymbol{Y}|$$

---

since the "centre dot" symbol is the preferred way to represent multiplication in those situations where juxtaposition causes confusion.

We can take the product of any finite number of sets (and sometimes of an infinite number of sets, although there are complications in doing so). Firstly, we extend the idea of an ordered pair. There also exist ordered triples, ordered quadruples, and so on.

Suppose that we have $\boldsymbol{n}$ sets, called $\boldsymbol{X_1, X_2, \ldots, X_n}$. Let $\boldsymbol{x_1 \in X_1}$, $\boldsymbol{x_2 \in X_2}$, $\ldots$, $\boldsymbol{x_n \in X_n}$. Then the object

$$\boldsymbol{(x_1, x_2, \ldots, x_n)}$$

is called an *ordered* $\boldsymbol{n}$*-tuple*. The name comes by generalising the idea of a quintuple, a sextuple, and so on—a quintuple can also be called a 5-tuple, a sextuple is a 6-tuple, et cetera.

The set of all such $\boldsymbol{n}$-tuples is denoted by

$$\boldsymbol{X_1 \times X_2 \times \cdots \times X_n}$$

and is called the *Cartesian product* of the sets.

Of particular importance is the situation in which all of the sets are the same set, $\boldsymbol{X}$ say. Then the Cartesian product is:

$$\underbrace{\boldsymbol{X \times X \times \cdots \times X}}_{\boldsymbol{n}\ \textbf{times}}$$

But this is cumbersome, and instead we

write:

$$X^n$$

So

$$X^n = \{(x_1, x_2, \ldots, x_n) : x_1, x_2, \ldots, x_n \in X\}.$$

This is the set of all ordered n-tuples from X. Note that:

$$|X^n| = |X|^n$$

Example:
The set $\mathbb{R}^3$ consists of ordered triples of the form (x, y, z). It is often used to describe the location of a point, relative to a system of three coordinate axes meeting at a fixed point called the "origin" of the system. The coordinates x, y and z give the distances one must travel in each of the three mutually perpendicular directions to get from the origin to the point.

End of example

Worked Example:
Consider the binary set $\mathbb{B} = \{0, 1\}$. We get

$$\mathbb{B}^2 = \{(0,0), (0,1), (1,0), (1,1)\}$$

while $\mathbb{B}^3$ has as its members these 8 ordered triples:

$$(0,0,0), (0,0,1), (0,1,0), (0,1,1), (1,0,0), (1,0,1), (1,1,0), (1,1,1)$$

End of worked example

The sets given in the last example play a vital role in the structure of Boolean algebras, as we shall see later.

Formative Assessment

You should now do as many practice exercises as necessary to establish that you can correctly determine the Cartesian products of sets.

End of formative assessment

Lecture 4

This lecture provides an introduction to:

- **binary operation on sets**
- **Boolean-algebraic properties of sets**

You should master the material contained in this lecture before moving on to the next lecture.

Operations on Sets

Before defining the most important operations on sets, we need to mention this relationship between equality and containment: if $\boldsymbol{X}$ and $\boldsymbol{Y}$ are two sets, then

$\boldsymbol{X = Y}$ if and only if

$$(\boldsymbol{X \subseteq Y} \text{ and } \boldsymbol{Y \subseteq X}).$$

This follows directly from the definition of containment. It is analogous to the fact that, for any two real numbers $\boldsymbol{a}$ and $\boldsymbol{b}$,

$\boldsymbol{a = b}$ if and only if ($\boldsymbol{a \leq b}$ and $\boldsymbol{b \leq a}$).

Let $\boldsymbol{X}$ and $\boldsymbol{Y}$ be sets. Their *union* $\boldsymbol{X \cup Y}$ is the set of all elements that are in $\boldsymbol{X}$ or $\boldsymbol{Y}$ (including those that are in both). So

$$\boldsymbol{X \cup Y = \{x : x \in X} \text{ or } \boldsymbol{x \in Y\}}$$

where the word "or" is being used here in an *inclusive* sense. The union is illustrated in the top diagram in Figure 1.

The *intersection* of $\boldsymbol{X}$ and $\boldsymbol{Y}$ is the set of all elements that they have in common. It is denoted by:

$$\boldsymbol{X \cap Y}$$

So

$$\boldsymbol{X \cap Y = \{x : x \in X} \text{ and } \boldsymbol{x \in Y\}}.$$

The intersection is illustrated in the bottom diagram in Figure 1.

Note that

$$\boldsymbol{X \cap Y \subseteq X \subseteq X \cup Y}$$

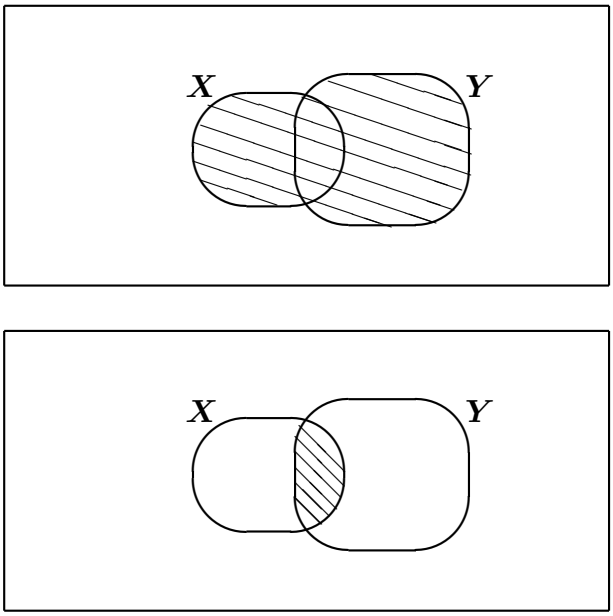

Figure 1: The union (top) and intersection (bottom) of two sets.

and that a similar statement is true if we replace the $\boldsymbol{X}$ in the middle by $\boldsymbol{Y}$. These containments follow directly from the definitions.

Worked Example:
We obtain the union and intersection of the two sets $\boldsymbol{X} = \{\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}, \boldsymbol{d}, \boldsymbol{e}\}$ and $\boldsymbol{Y} = \{\boldsymbol{a}, \boldsymbol{c}, \boldsymbol{e}, \boldsymbol{f}\}$. To get their union, we can begin by listing all the elements of $\boldsymbol{X}$:

$$\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}, \boldsymbol{d}, \boldsymbol{e}$$

Now we look at each element of $\boldsymbol{Y}$ in turn, and add it to the list if it hasn't already been mentioned. We find that $\boldsymbol{a}$, $\boldsymbol{c}$ and $\boldsymbol{e}$ are already mentioned in the list we have started creating, but $\boldsymbol{f}$ isn't. So we add $\boldsymbol{f}$ to the list, which is then:

$$\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}, \boldsymbol{d}, \boldsymbol{e}, \boldsymbol{f}$$

Making a set out of this gives us the union:

$$\boldsymbol{X} \cup \boldsymbol{Y} = \{\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}, \boldsymbol{d}, \boldsymbol{e}, \boldsymbol{f}\}$$

To get the intersection, we look at each element of $\boldsymbol{X}$ in turn. We see if that element is also in $\boldsymbol{Y}$. If so, we write it down. This creates a list of the elements in the intersection. The list is

$$\boldsymbol{a}, \boldsymbol{c}, \boldsymbol{e}$$

and making a set out of it gives us the intersection

$$\boldsymbol{X} \cap \boldsymbol{Y} = \{\boldsymbol{a}, \boldsymbol{c}, \boldsymbol{e}\}.$$

End of worked example

Remark

A *lemma* is a "little theorem", useful in proving other theorems.

Lemma:
For any two sets $\boldsymbol{X}$ and $\boldsymbol{Y}$,

$$|\boldsymbol{X} \cup \boldsymbol{Y}| = |\boldsymbol{X}| + |\boldsymbol{Y}| - |\boldsymbol{X} \cap \boldsymbol{Y}|.$$

End of lemma

Proof:

If we count the elements in $\boldsymbol{X}$, and then count the elements in $\boldsymbol{Y}$, and add the totals, we get $|\boldsymbol{X}|+|\boldsymbol{Y}|$. But the elements in the intersection have been counted twice, once when they occurred as elements of $\boldsymbol{X}$ and then again when they occurred as elements of $\boldsymbol{Y}$. So we subtract **1** for each element in the intersection, which means that we have to subtract $|\boldsymbol{X} \cap \boldsymbol{Y}|$ from the sum $|\boldsymbol{X}| + |\boldsymbol{Y}|$ to get the exact number of elements in the union.

End of proof

Worked Example:
We demonstrate that the Lemma is true

for the sets given in the last worked example. So $\boldsymbol{X = \{a, b, c, d, e\}}$ and $\boldsymbol{Y = \{a, c, e, f\}}$. We found that

$$\boldsymbol{X \cup Y = \{a, b, c, d, e, f\}}$$

and

$$\boldsymbol{X \cap Y = \{a, c, e\}.}$$

We want to verify that

$$\boldsymbol{|X \cup Y| = |X| + |Y| - |X \cap Y|}$$

in this case. Here is the verification:

$$\begin{array}{rcl} \boldsymbol{LHS} & = & \boldsymbol{|X \cup Y|} \\ & = & \boldsymbol{|\{a, b, c, d, e, f\}|} \\ & = & \boldsymbol{6} \\ \boldsymbol{RHS} & = & \boldsymbol{|X| + |Y| - |X \cap Y|} \\ & = & \boldsymbol{5 + 4 - 3} \\ & = & \boldsymbol{6} \\ & = & \boldsymbol{LHS} \text{ as required.} \end{array}$$

So the Lemma holds for this particular example.

End of worked example

Remark

Of course, because we have *proved* the lemma, it holds for **every** example. So there is really no need to verify it for individual examples.

If two sets don't intersect, so that $\boldsymbol{X \cap Y = \varnothing}$, they are said to be *disjoint* and their union is also called their *disjoint union*. It is denoted by:

$$\boldsymbol{X \dot{\cup} Y}$$

Remark

A *corollary* is a result that follows from a previous result.

The following corollary is an immediate consequence of the last theorem.

Corollary:
Two sets $\boldsymbol{X}$ and $\boldsymbol{Y}$ are disjoint if and only if

$$|\boldsymbol{X} \cup \boldsymbol{Y}| = |\boldsymbol{X}| + |\boldsymbol{Y}|.$$

End of corollary

Notice in Figure 1 that the union of $\boldsymbol{X}$ and $\boldsymbol{Y}$ is made up of 3 disjoint regions. The middle region is the intersection of $\boldsymbol{X}$ and $\boldsymbol{Y}$. But how do we describe the other two regions?

The right-hand region inside the union consists of those elements in $\boldsymbol{Y}$ which are not in $\boldsymbol{X}$. This set is called the *complement of* $\boldsymbol{X}$ *relative to* $\boldsymbol{Y}$. It is denoted by

$$\boldsymbol{Y} \setminus \boldsymbol{X}$$

although some text books use the alternative notation

$$\boldsymbol{Y} - \boldsymbol{X}.$$

So

$$\boldsymbol{Y} \setminus \boldsymbol{X} = \{\boldsymbol{x} \in \boldsymbol{Y} : \boldsymbol{x} \notin \boldsymbol{X}\}.$$

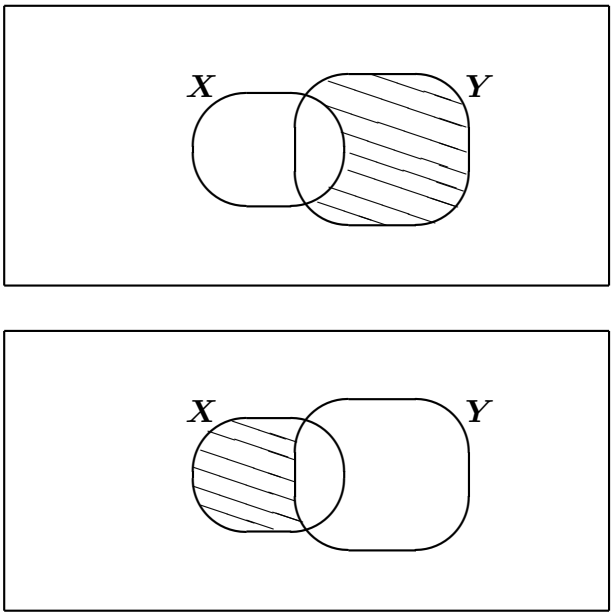

Figure 2: Relative complements $\boldsymbol{Y} \setminus \boldsymbol{X}$ (top) and $\boldsymbol{X} \setminus \boldsymbol{Y}$ (bottom).

This set is shaded in the top diagram of Figure 2.

Similarly, the left-hand region inside the union is

$$\boldsymbol{X \setminus Y}$$

which consists of those elements of $\boldsymbol{X}$ which are not in $\boldsymbol{Y}$. This set is the complement of $\boldsymbol{Y}$ relative to $\boldsymbol{X}$. It is shaded in the bottom diagram of Figure 2.

Notice that combining $\boldsymbol{X}$ with $\boldsymbol{Y \setminus X}$ gives us the union of $\boldsymbol{X}$ and $\boldsymbol{Y}$. It is because of this that we use the term "relative complement" to describe $\boldsymbol{Y \setminus X}$. The disjoint sets $\boldsymbol{X}$ and $\boldsymbol{Y \setminus X}$ **complement** each other, in the sense that together they make up the whole of $\boldsymbol{X \cup Y}$.

That is,

$$\boldsymbol{X \cup Y = X \ \dot\cup \ (Y \setminus X).}$$

Note also that

$$\boldsymbol{X = (X \cap Y) \ \dot\cup \ (X \setminus Y).}$$

If follows that

$$\boldsymbol{(X \setminus Y) = X \setminus (X \cap Y).}$$

The use of the "back-slash" symbol here is very evocative, especially in the right-hand side. To get $\boldsymbol{X \setminus Y}$ we **slash** from $\boldsymbol{X}$ its subset $\boldsymbol{X \cap Y}$.

Worked Example:

We find the relative complements $\boldsymbol{X \setminus Y}$ and $\boldsymbol{Y \setminus X}$ for the sets $\boldsymbol{X}$ and $\boldsymbol{Y}$ given in the previous worked examples.

To get $\boldsymbol{X \setminus Y}$ we look at each element of $\boldsymbol{X}$ in turn. We see if that element is also in $\boldsymbol{Y}$. If not, we write it down. This creates a list of the elements we want. The list is

$$\boldsymbol{b, d}$$

and making a set out of it gives us the relative complement

$$\boldsymbol{X \setminus Y = \{b, d\}.}$$

The statement

$$\boldsymbol{Y \setminus X = \{f\}}$$

can be derived similarly.

End of worked example

The easiest way to work with complements is to define them in terms of a "universal set". Consider again Figures 1–2. In each case there is a rectangular box enclosing the sets. This can itself be regarded as a set—the *universal set*. It represents the universe in which all of the sets currently under study occur. We often denote it by $\boldsymbol{\mathcal{U}}$.

The set $\boldsymbol{\mathcal{U}}$ is not universal in an **absolute** sense. It only needs to be big enough to contain as subsets all of the sets that are of interest to us in a particular situation.

So the universal set may vary from example to example.

The *complement* $\overline{\boldsymbol{X}}$ of the set $\boldsymbol{X}$ is defined to be its complement *relative to the universal set* $\boldsymbol{\mathcal{U}}$. That is,

$$\overline{\boldsymbol{X}} = \boldsymbol{\mathcal{U}} \setminus \boldsymbol{X}.$$

This is depicted in the top diagram of Figure 3. Clearly, for every element $\boldsymbol{x}$ of the universal set $\boldsymbol{\mathcal{U}}$,

$$\boldsymbol{x} \in \overline{\boldsymbol{X}} \text{ if and only if } \boldsymbol{x} \notin \boldsymbol{X}.$$

When there is a universal set $\boldsymbol{\mathcal{U}}$ available, all of the relative complements can be defined in terms of complements relative to $\boldsymbol{\mathcal{U}}$. The description is given in the following lemma.

Lemma:
For any two subsets $\boldsymbol{X}$ and $\boldsymbol{Y}$ of $\boldsymbol{\mathcal{U}}$,

$$\boldsymbol{X} \setminus \boldsymbol{Y} = \boldsymbol{X} \cap \overline{\boldsymbol{Y}}.$$

End of lemma

Proof:
As mentioned at the start of this lecture, two sets are equal if each is contained in the other. So to prove the lemma we will firstly show that $\boldsymbol{LHS} \subseteq \boldsymbol{RHS}$ and then that $\boldsymbol{RHS} \subseteq \boldsymbol{LHS}$.

Let $\boldsymbol{x} \in \boldsymbol{X} \setminus \boldsymbol{Y}$. Then $\boldsymbol{x} \in \boldsymbol{X}$ and $\boldsymbol{x} \notin \boldsymbol{Y}$. Since $\boldsymbol{x} \notin \boldsymbol{Y}$, we know that $\boldsymbol{x} \in \overline{\boldsymbol{Y}}$. So we

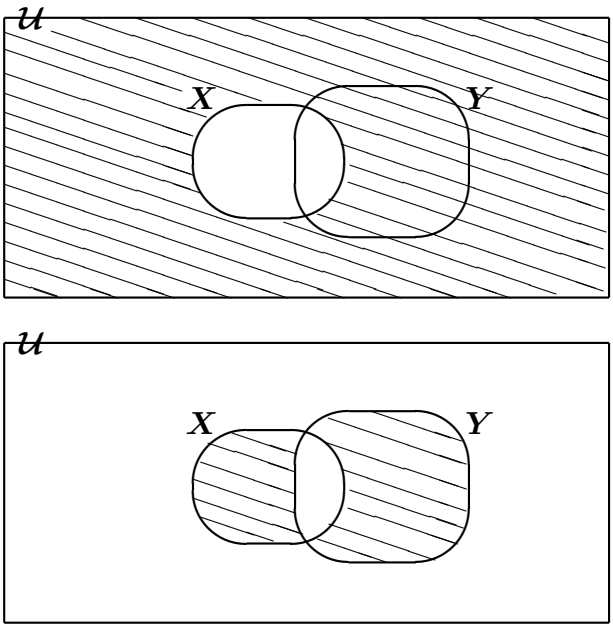

Figure 3: The complement (of $\boldsymbol{X}$, top) and symmetric difference (bottom).

know that $\boldsymbol{x} \in \boldsymbol{X}$ and $\boldsymbol{x} \in \overline{\boldsymbol{Y}}$. It follows immediately (from the definition of "intersection") that $\boldsymbol{x} \in \boldsymbol{X} \cap \overline{\boldsymbol{Y}}$. So we have proved that

$$\boldsymbol{X \setminus Y \subseteq X \cap \overline{Y}.}$$

Let $\boldsymbol{x} \in \boldsymbol{X} \cap \overline{\boldsymbol{Y}}$. Then $\boldsymbol{x} \in \boldsymbol{X}$ and $\boldsymbol{x} \in \overline{\boldsymbol{Y}}$. Since $\boldsymbol{x} \in \overline{\boldsymbol{Y}}$, we know that $\boldsymbol{x} \notin \boldsymbol{Y}$. So $\boldsymbol{x} \in \boldsymbol{X}$ and $\boldsymbol{x} \notin \boldsymbol{Y}$. It follows immediately (from the definition of "relative complement") that $\boldsymbol{x} \in \boldsymbol{X} \setminus \boldsymbol{Y}$. So we have proved that

$$\boldsymbol{X \cap \overline{Y} \subseteq X \setminus Y.}$$

The two containments give us the equality we need.

End of proof

Worked Example:
If the universal set is

$$\boldsymbol{\mathcal{U} = \{1, 2, \ldots, 10\},}$$

$\boldsymbol{X}$ consists of the numbers **1** to **5**, and $\boldsymbol{Y}$ consists of the numbers **3**, **4**, **5** and **6**, then we can find the complement of each of $\boldsymbol{X}$, $\boldsymbol{Y}$ and $\boldsymbol{X \cup Y}$. Firstly we have:

$\boldsymbol{\overline{X} = \{6, 7, 8, 9, 10\}}$
$\boldsymbol{\overline{Y} = \{1, 2, 7, 8, 9, 10\}}$

Next, it's easy to see that the union has **1**, **2**, **3**, **4**, **5** and **6** in it. So its complement

$$\boldsymbol{\overline{X \cup Y}}$$

consists of the numbers **7**, **8**, **9** and **10**.

End of worked example

Remark

It might be expected that

$$\overline{\boldsymbol{X} \cup \boldsymbol{Y}}$$

would be equal to

$$\overline{\boldsymbol{X}} \cup \overline{\boldsymbol{Y}}$$

but we can see that in this last worked example it isn't so. In fact, it is the *intersection* of $\overline{\boldsymbol{X}}$ and $\overline{\boldsymbol{Y}}$ which equals the complement of their union. This is **always** true, for any two sets $\boldsymbol{X}$ and $\boldsymbol{Y}$ contained in some universal set $\boldsymbol{\mathcal{U}}$; and furthermore, interchanging the words "intersection" and "union" in the last sentence gives another statement which is always true. These two statements are called *de Morgan's laws* for the union and intersection of sets. We will meet these laws in a more general setting when we study Boolean Algebra.

Sometimes we give a special name to that subset of $\boldsymbol{X} \cup \boldsymbol{Y}$ which excludes the intersection $\boldsymbol{X} \cap \boldsymbol{Y}$. It is called the *symmetric difference* of $\boldsymbol{X}$ and $\boldsymbol{Y}$ and is denoted by:

$$\boldsymbol{X} \triangle \boldsymbol{Y}$$

So

$$\boldsymbol{X} \triangle \boldsymbol{Y} = (\boldsymbol{X} \cup \boldsymbol{Y}) \setminus (\boldsymbol{X} \cap \boldsymbol{Y}).$$

The symmetric difference is illustrated in the bottom diagram of Figure 3. Clearly, we can also say that

$$\boldsymbol{X \triangle Y = (X \setminus Y) \ \dot\cup \ (Y \setminus X).}$$

Worked Example:
We return to the first worked example, where $\boldsymbol{X = \{a, b, c, d, e\}}$ and $\boldsymbol{Y = \{a, c, e, f\}}$. We found their union to be

$$\boldsymbol{X \cup Y = \{a, b, c, d, e, f\}}$$

and their intersection to be

$$\boldsymbol{X \cap Y = \{a, c, e\}.}$$

Slashing the intersection from the union gives

$$\boldsymbol{X \triangle Y = \{b, d, f\}.}$$

But we have also seen that the symmetric difference is the (disjoint) union of the relative complements. Let's check to see if that gives us the same answer. In another worked example we found that

$$\boldsymbol{X \setminus Y = \{b, d\}}$$

and that

$$\boldsymbol{Y \setminus X = \{f\}.}$$

They are certainly disjoint, and their union is the set

$$\boldsymbol{\{b, d, f\}}$$

as required.

End of worked example

It is now an appropriate point in our study to gather together some of the most important properties displayed by a family of sets for which a universal set exists.

Let $\boldsymbol{A}$, $\boldsymbol{B}$ and $\boldsymbol{C}$ be subsets of some universal set $\boldsymbol{\mathcal{U}}$. (Note that the symbol $\boldsymbol{B}$ here is representing some arbitrary set. It is not necessarily the binary set $\mathbb{B}$.) Then the following properties hold:

$$\boldsymbol{A} \cup \boldsymbol{B} = \boldsymbol{B} \cup \boldsymbol{A}$$
$$\boldsymbol{A} \cap \boldsymbol{B} = \boldsymbol{B} \cap \boldsymbol{A}$$
$$(\boldsymbol{A} \cup \boldsymbol{B}) \cup \boldsymbol{C} = \boldsymbol{A} \cup (\boldsymbol{B} \cup \boldsymbol{C})$$
$$(\boldsymbol{A} \cap \boldsymbol{B}) \cap \boldsymbol{C} = \boldsymbol{A} \cap (\boldsymbol{B} \cap \boldsymbol{C})$$
$$\boldsymbol{A} \cap (\boldsymbol{B} \cup \boldsymbol{C}) = (\boldsymbol{A} \cap \boldsymbol{B}) \cup (\boldsymbol{A} \cap \boldsymbol{C})$$
$$\boldsymbol{A} \cup (\boldsymbol{B} \cap \boldsymbol{C}) = (\boldsymbol{A} \cup \boldsymbol{B}) \cap (\boldsymbol{A} \cup \boldsymbol{C})$$
$$\boldsymbol{A} \cup \boldsymbol{\varnothing} = \boldsymbol{A}$$
$$\boldsymbol{A} \cap \boldsymbol{\mathcal{U}} = \boldsymbol{A}$$
$$\boldsymbol{A} \cup \overline{\boldsymbol{A}} = \boldsymbol{\mathcal{U}}$$
$$\boldsymbol{A} \cap \overline{\boldsymbol{A}} = \boldsymbol{\varnothing}$$

The first two properties are called the *commutative* laws. The next two are the *associative* laws. Then we have the two *distributive* laws. After them come the two *identity* laws. The last two are the *complement* laws.

Because of these properties, the power set

$\mathcal{P}(\mathcal{U})$ (consisting of all the subsets of $\mathcal{U}$) is an example of the mathematical system known as a *Boolean algebra*. More particularly, it is a *Boolean algebra of sets*. We will revisit this idea in a later lecture.

Formative Assessment

You should now do as many practice exercises as necessary to establish that you can correctly find unions, intersections, complements and symmetric differences of sets.

End of formative assessment

Lecture 5

This lecture provides an introduction to:

- **functions between sets**
- **properties of functions between finite sets**
- **infinite sets**

You should master the material contained in this lecture before moving on to the next lecture.

Functions between Sets

Before beginning our study of functions, we need to mention a feature of sets which may not yet have been made explicit.

It is important to understand that a set is a *collection* of objects rather than a "list" of objects. What this distinction means is that *the order in which the elements in a set are listed is of no importance*, in the sense that changing the order does not change the set.

So we can say that

$$\{1, 2, 3, 4\} = \{4, 2, 3, 1\}$$

and so on.

Furthermore, elements can be mentioned more than once without the set being changed. For example,

$$\{1, 2, 1, 3, 1, 4, 2\} = \{1, 2, 3, 4\}$$

although the left-hand set might at first glance appear to have more elements in it. The right-hand set is just a simpler (and usually preferable) way to represent the left-hand set.

A *function* is a procedure for systematically changing the individual elements of a set into elements of another set.

Let $\boldsymbol{X}$ and $\boldsymbol{Y}$ be sets. A function $\boldsymbol{f}$ from $\boldsymbol{X}$ to $\boldsymbol{Y}$ changes each element of $\boldsymbol{X}$ into an element of $\boldsymbol{Y}$. We call $\boldsymbol{X}$ the *domain* and $\boldsymbol{Y}$ the *codomain* of the function. Let

$\boldsymbol{x} \in \boldsymbol{X}$, and let $\boldsymbol{y}$ be the element of $\boldsymbol{Y}$ into which $\boldsymbol{x}$ is changed. We call $\boldsymbol{y}$ the *image* of $\boldsymbol{x}$, under the action of $\boldsymbol{f}$, and denote it by $\boldsymbol{f(x)}$ (which is read as "$\boldsymbol{f}$ *of* $\boldsymbol{x}$"). So $\boldsymbol{f(x) = y}$. We also say that $\boldsymbol{x}$ *is mapped to* $\boldsymbol{y}$.

Note that it is allowable for $\boldsymbol{X}$ and $\boldsymbol{Y}$ to be one and the same set.

A function is also called a *mapping* or a *map*. We say that $\boldsymbol{f}$ *maps the set* $\boldsymbol{X}$ *into the set* $\boldsymbol{Y}$.

The display

$$\begin{array}{rcl} \boldsymbol{f : X} & \boldsymbol{\rightarrow} & \boldsymbol{Y} \\ \boldsymbol{x} & \boldsymbol{\mapsto} & \boldsymbol{f(x)} \end{array}$$

means that $\boldsymbol{f}$ is a function from the set $\boldsymbol{X}$ to the set $\boldsymbol{Y}$, and that $\boldsymbol{f}$ maps a typical element $\boldsymbol{x}$ of $\boldsymbol{X}$ to the element $\boldsymbol{f(x)}$ in $\boldsymbol{Y}$.

Note that an ordinary arrow points from the domain to the codomain, but the arrow that points from the typical element to its image has a short T-bar on the end of its tail.

Worked Example:
Let $\boldsymbol{X = \{-2, -1, 0, 1, 2\}}$ and let $\boldsymbol{Y = \{0, 1, 2, 3, 4, 5\}}$. Suppose that a function $\boldsymbol{f : X \rightarrow Y}$ is defined by $\boldsymbol{f(x) = x^2}$. What are the images of the elements in $\boldsymbol{X}$?

We get

$$\begin{array}{rcrcl} f(-2) & = & (-2)^2 & = & 4 \\ f(-1) & = & (-1)^2 & = & 1 \\ f(0) & = & (0)^2 & = & 0 \\ f(1) & = & (1)^2 & = & 1 \\ f(2) & = & (2)^2 & = & 4 \end{array}$$

so that the three numbers 0, 1 and 4 appear as images.

End of worked example

This worked example illustrates several important features. Firstly, although each element of X has only one image (which is the main requirement for f to be a function), several elements of Y occurred as the image of more than one element. Secondly, some elements of Y did not occur as images of elements from X. These observations lead to some further definitions.

If $f : X \to Y$ is a function, then the set of all images of elements from the domain X is a subset of the codomain Y, called the *range* of the function f. It is denoted by $f(X)$. So we can write that

$$\begin{array}{rcl} f(X) & = & \{f(x) \in Y : x \in X\} \\ & = & \text{the } \textbf{range} \text{ of } f. \end{array}$$

When the range of f is the *entire* codomain Y, we say that f maps X *onto* Y. We also call f an *onto* function. The more sophisticated name is: *surjection*. So a function $f : X \to Y$ is *surjective* if and only if $f(X) = Y$.

In the last worked example, the codomain Y is the set $\{0, 1, 2, 3, 4, 5\}$ but the range is $\{0, 1, 4\}$. So this function is not surjective.

Also in that example there were some elements of the range which occurred more than once as images of domain elements. For example, $f(-2) = 4$ and also $f(2) = 4$ so that 4 is the image of two different elements, -2 and 2. Such a function is called "many-to-one", because it's possible to find two distinct elements that map to the one image.

When no such combination can be found, a function is said to be *one-to-one* or to be an *injection*. So a function is *injective* if and only if

$$x_1 \neq x_2 \Rightarrow f(x_1) \neq f(x_2).$$

That is, different domain elements always have different images.

For any function f, if $f(x) = y$ then as well as saying that y is the image of x we also say that x is *a pre-image* of y. When y has only one pre-image (and in particular when f is one-to-one) we can speak of *the pre-image* of y.

Note also that a function is onto if and only if every element in the codomain has at least one pre-image.

Worked Example:
Consider the function f, defined as follows:

$$\begin{aligned} f: \ & \mathbb{N} \to \mathbb{N} \\ & n \mapsto n+1 \end{aligned}$$

This function maps 1 to 2, 2 to 3, 3 to 4, and so on. It is easily seen to be one-to-one. For, suppose that two elements n_1 and n_2 had the same image. So $f(n_1) = f(n_2)$. That is:

$$n_1 + 1 = n_2 + 1$$

Subtracting 1 from both sides gives $n_1 = n_2$. So the only time that "two" elements n_1 and n_2 can both have the same image is if they are in fact one and the same element. This shows that the function is one-to-one.

But clearly the function is not onto, as the number 1 has no pre-image.

End of worked example

A mapping that is both one-to-one and onto, being both injective and surjective, is called a *bijection* or a *one-to-one correspondence*.

Worked Example:
Consider the function f, defined as follows:

$$\begin{aligned} f: \ & \mathbb{Z} \to \mathbb{Z} \\ & n \mapsto n+1 \end{aligned}$$

This function has the same effect as that in the last worked example, but now it is defined for **all** integers rather than just for the positive integers.

The explanation in the last worked example shows why this function is one-to-one. It is also onto, because the pre-image of any element n of $\mathbb{Z}$ is the element $n - 1$ which is also in $\mathbb{Z}$. So this function provides us with an example of a one-to-one correspondence.

End of worked example

Formative Assessment

You should now do as many practice exercises as necessary to establish that you can correctly determine if a given function is one-to-one or onto.

End of formative assessment

There are some very important links between the relative cardinalities of two sets and the existence of different kinds of functions between those two sets. For finite sets, these links are developed in the next section. Functions involving infinite sets can sometimes exhibit unexpected behaviour. So for a while we will concentrate on functions between finite sets.

Functions between Finite Sets

Throughout this section all sets will be regarded as finite, unless otherwise specified.

Lemma:

If $\boldsymbol{X \subseteq Y}$ then $\boldsymbol{|X| \leq |Y|}$.

End of lemma

Proof:
If $\boldsymbol{X = Y}$ then clearly $\boldsymbol{|X| = |Y|}$. If instead $\boldsymbol{X \subset Y}$ then we can write $\boldsymbol{X = \{x_1, \ldots, x_m\}}$ and $\boldsymbol{Y = \{x_1, \ldots, x_n\}}$ where $\boldsymbol{m < n}$. So $\boldsymbol{|X| = m}$ and $\boldsymbol{|Y| = n}$, whence $\boldsymbol{|X| < |Y|}$. In either case we have $\boldsymbol{|X| \leq |Y|}$.

End of proof

Theorem:
If a function $\boldsymbol{f}$ is defined on a set $\boldsymbol{X}$, then

$$\boldsymbol{|f(X)| \leq |X|.}$$

End of theorem

Proof:
Suppose that $\boldsymbol{X = \{x_1, \ldots, x_m\}}$. Then the range of $\boldsymbol{f}$ is

$$\boldsymbol{f(X) = \{f(x_1), \ldots, f(x_m)\}.}$$

If $\boldsymbol{f}$ is one-to-one, then all of these elements are distinct, so that there are $\boldsymbol{m}$ of them. In this case

$$\boldsymbol{|f(X)| = m = |X|.}$$

If $\boldsymbol{f}$ is not one-to-one, then there is some repetition within the set

$$\boldsymbol{\{f(x_1), \ldots, f(x_m)\}.}$$

Thus its cardinality, which is $|f(X)|$, is less than m. So in both cases $|f(X)| \leq m$. That is, $|f(X)| \leq |X|$ as required.

End of proof

Theorem:
Let X and Y be finite sets. Then $|X| \leq |Y|$ if and only if there exists a one-to-one function from X into Y.

End of theorem

Proof:
Assume that $|X| \leq |Y|$. Then we can write

$$X = \{x_1, \ldots, x_m\}$$

and

$$Y = \{y_1, \ldots, y_n\}$$

where $m \leq n$. Define a mapping $f : X \to Y$ by $f(x_i) = y_i$ for each $i = 1, \ldots, m$. Since all the y_i are distinct, this is one-to-one.

Conversely, assume that there exists a one-to-one function f from X into Y. From the proof of the last theorem we know that

$$|X| = |f(X)|.$$

Since the range $f(X)$ is contained in the codomain Y, we know from the lemma that $|f(X)| \leq |Y|$. Putting these two facts together gives us that $|X| \leq |Y|$ as required.

End of proof

Theorem:
Let $\boldsymbol{X}$ and $\boldsymbol{Y}$ be finite sets. Then $|\boldsymbol{X}| \geq |\boldsymbol{Y}|$ if and only if there exists a function from $\boldsymbol{X}$ onto $\boldsymbol{Y}$.

End of theorem

Proof:
Assume that $|\boldsymbol{X}| \geq |\boldsymbol{Y}|$. Then we can write

$$\boldsymbol{X} = \{\boldsymbol{x_1}, \ldots, \boldsymbol{x_m}\}$$

and

$$\boldsymbol{Y} = \{\boldsymbol{y_1}, \ldots, \boldsymbol{y_n}\}$$

where $\boldsymbol{m} \geq \boldsymbol{n}$. Define a mapping $\boldsymbol{f}$: $\boldsymbol{X} \rightarrow \boldsymbol{Y}$ by $\boldsymbol{f(x_i)} = \boldsymbol{y_i}$ for each $\boldsymbol{i} = \boldsymbol{1, \ldots, n}$. If there are still some elements in $\boldsymbol{X}$ with no images (specifically $\boldsymbol{x_n + 1}$ to $\boldsymbol{x_m}$), map them all to $\boldsymbol{y_1}$ (or to any other element of $\boldsymbol{Y}$). Since every element of $\boldsymbol{Y}$ is the image of at least one of the $\boldsymbol{x_i}$, the function $\boldsymbol{f}$ is onto.

Conversely, assume that there exists a function $\boldsymbol{f}$ from $\boldsymbol{X}$ onto $\boldsymbol{Y}$. Since $\boldsymbol{f}$ is onto, the range and the codomain coincide; that is, $\boldsymbol{f(X)} = \boldsymbol{Y}$. Therefore $|\boldsymbol{f(X)}| = |\boldsymbol{Y}|$. As always, $|\boldsymbol{f(X)}| \leq |\boldsymbol{X}|$. Replacing $|\boldsymbol{f(X)}|$ by $|\boldsymbol{Y}|$ in this inequality gives $|\boldsymbol{Y}| \leq |\boldsymbol{X}|$ as required.

End of proof

Theorem:
Let $\boldsymbol{X}$ and $\boldsymbol{Y}$ be finite sets. Then $|\boldsymbol{X}| = |\boldsymbol{Y}|$ if and only if there exists a one-to-one function from $\boldsymbol{X}$ onto $\boldsymbol{Y}$.

End of theorem

Proof:
This follows immediately from the two previous theorems.

End of proof

Worked Example:
Let $\boldsymbol{X} = \{\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}, \boldsymbol{d}\}$ and $\boldsymbol{Y} = \{\mathbf{1}, \mathbf{2}, \mathbf{3}\}$.

A function from $\boldsymbol{X}$ into $\boldsymbol{Y}$ must provide an image for each element from $\boldsymbol{X}$. To make the range as small as possible we could define a function $\boldsymbol{f}$ that maps every element of $\boldsymbol{X}$ to one element of $\boldsymbol{Y}$, by declaring that

$$\boldsymbol{f(a) = f(b) = f(c) = f(d) = 2}$$

for example. On the other hand, we could define a function $\boldsymbol{g}$ for which the range is as large as possible, by putting

$$\boldsymbol{g(a) = 1, g(b) = 2, g(c) = g(d) = 3}$$

for example. This function $\boldsymbol{g}$ is onto. It is impossible to define a one-to-one function from $\boldsymbol{X}$ into $\boldsymbol{Y}$, because we would need to

have at least 4 elements in $\boldsymbol{Y}$ to serve as the distinct images of the 4 elements from $\boldsymbol{X}$.

But a one-to-one function can be defined from $\boldsymbol{Y}$ into $\boldsymbol{X}$. For example, the function $\boldsymbol{h : Y \rightarrow X}$ defined by

$$\boldsymbol{h(1) = a, h(2) = b, h(3) = c}$$

is one-to-one but not onto. Because $\boldsymbol{X}$ and $\boldsymbol{Y}$ have different cardinalities, there is no function from one to the other (in either direction) which is both one-to-one and onto.

End of worked example

The last three theorems give conditions for the existence of certain kinds of functions. One question that arises is this. How many functions are there between two sets? This is addressed in the next theorem.

Theorem:
Let $\boldsymbol{X}$ and $\boldsymbol{Y}$ be finite sets. Then the number of different functions from $\boldsymbol{X}$ into $\boldsymbol{Y}$ is given by the formula

$$|\boldsymbol{Y}|^{|\boldsymbol{X}|}$$

which depends on the number of elements in $\boldsymbol{X}$ and the number of elements in $\boldsymbol{Y}$.

End of theorem

Proof:

Put $\boldsymbol{X = \{x_1, \ldots, x_m\}}$ and $\boldsymbol{Y = \{y_1, \ldots, y_n\}}$. Now the element $\boldsymbol{x_1}$ could have any of $\boldsymbol{y_1, \ldots, y_n}$ as its image. So there are $\boldsymbol{n}$ possibilities.

Similarly

$\boldsymbol{x_2}$ has $\boldsymbol{n}$ possible images,

$\vdots$

and $\boldsymbol{x_m}$ has $\boldsymbol{n}$ possible images.

So the total number of combinations is

$$\underbrace{\boldsymbol{n \cdot n \ldots n}}_{\boldsymbol{m \text{ times}}}$$

which equals $\boldsymbol{n^m}$.

End of proof

Because of this last result, the *set* of all possible functions from $\boldsymbol{X}$ to $\boldsymbol{Y}$ is sometimes denoted by:

$$\boldsymbol{Y^X}$$

We can then say that:

$$\boldsymbol{|Y^X| = |Y|^{|X|}}$$

Remark

It is possible, without too much difficulty, to develop formulae for the number of one-to-one functions and the number of onto functions between two sets. But before doing so we would need to spend some time studying the topic of "methods of counting". These ideas are developed further in

most standard books on Discrete Mathematics.

We now describe an important kind of function.

Definition:
Let $\boldsymbol{A}$ be a subset of a set $\boldsymbol{X}$. Then the *characteristic function* $\boldsymbol{\chi_A}$ is a function, from $\boldsymbol{X}$ into the binary set $\mathbb{B} = \boldsymbol{\{0, 1\}}$, defined as follows:

$$\boldsymbol{\chi_A}(\boldsymbol{x}) = \begin{cases} \mathbf{1} \text{ if } \boldsymbol{x} \in \boldsymbol{A} \\ \mathbf{0} \text{ otherwise} \end{cases}$$

So the function has value $\mathbf{1}$ when evaluated at the elements of the subset $\boldsymbol{A}$, but value $\mathbf{0}$ when evaluated at the elements of $\boldsymbol{X}$ which are not in $\boldsymbol{A}$.

End of definition

Note that the symbol $\boldsymbol{\chi}$ is the lower-case Greek letter *chi*.

Worked Example:
Let $\boldsymbol{X} = \boldsymbol{\{a, b, c, d, e\}}$ and let $\boldsymbol{A}$ be the subset of $\boldsymbol{X}$ consisting of $\boldsymbol{a}$, $\boldsymbol{c}$ and $\boldsymbol{e}$. Then the characteristic function is defined by:

$$\begin{aligned}\chi_A(a) &= 1\\ \chi_A(b) &= 0\\ \chi_A(c) &= 1\\ \chi_A(d) &= 0\\ \chi_A(e) &= 1\end{aligned}$$

The elements a, c and e are all mapped to 1, because they are in the subset A. The other elements are mapped to 0.

End of worked example

Lemma:
Every function from a set X to the set $\mathbb{B}$ is the characteristic function of some subset of X.

End of lemma

Proof:
Let f be a function from X into $\mathbb{B}$. Let A be the set of all pre-images, under the function f, of the element 1 of $\mathbb{B}$. It is clear that f is precisely the characteristic function χ_A.

End of proof

Corollary:
The number of subsets of a finite set X is

$$2^{|X|}$$

so that the cardinality of the power set is

always a power of 2.

End of corollary

Proof:
It follows from the lemma that the number of subsets of $\boldsymbol{X}$ is precisely the number of functions from $\boldsymbol{X}$ into $\mathbb{B}$. But by a previous theorem the number of functions from $\boldsymbol{X}$ into $\mathbb{B}$ is:

$$|\mathbb{B}|^{|\boldsymbol{X}|}$$

Since $\mathbb{B}$ is a $\mathbf{2}$-element set, this formula can be expressed more simply as:

$$\mathbf{2}^{|\boldsymbol{X}|}$$

This formula was demonstrated in a previous lecture, but as shown above we now have the machinery necessary to prove that it is correct.

End of proof

We now explain a useful way to represent functions between small finite sets.

Let $\boldsymbol{f} : \boldsymbol{X} \rightarrow \boldsymbol{Y}$ be a function. Suppose that

$$\boldsymbol{X} = \{\boldsymbol{a_1}, \ldots, \boldsymbol{a_m}\}$$

and

$$\boldsymbol{Y} = \{\boldsymbol{b_1}, \ldots, \boldsymbol{b_n}\}.$$

The images of

$$\boldsymbol{a_1}, \ldots, \boldsymbol{a_m}$$

are

$$\boldsymbol{f}(\boldsymbol{a_1}), \ldots, \boldsymbol{f}(\boldsymbol{a_m})$$

respectively. Suppose we call these

$$c_1, \ldots, c_m$$

where each c_i is in Y but we don't necessarily know which elements of Y they are.

Now we can write down the set of ordered pairs

$$\{(a_1, c_1), \ldots, (a_m, c_m)\}$$

where in each ordered pair the first coordinate is from the domain X and the second coordinate is the image of the first.

Then, provided also that we know the codomain Y, this set of ordered pairs completely determines the function f, and vice versa. For, we know everything about f if we know where it maps each element of the domain.

For this reason, such a set of ordered pairs is often *equated* with the associated function. That is, we say that the function *equals* the set of ordered pairs:

$$f = \{(a_1, c_1), \ldots, (a_m, c_m)\}$$

From this point of view, *a function can be regarded as a special kind of subset of the Cartesian product* $X \times Y$. Specifically, it is a subset of $X \times Y$ in which each element of X appears as the first coordinate in exactly one of the ordered pairs.

More generally, a *relation* between the sets X and Y (either of which can be infinite)

is defined to be **any** subset of $\boldsymbol{X \times Y}$ whatsoever. So a function is a special kind of relation. Loosely speaking, relations can be regarded as "multivalued" mappings from subsets of the domain into the codomain.

The study of relations is an important topic in higher discrete mathematics, and is covered in many of the standard text books.

If a function has a small domain, then the function can be displayed by means of a two-column table. The first column shows the domain elements, and in the second column are the images of those elements.

Example:

The characteristic function described in the last worked example can be displayed as follows.

x	$\chi_A(x)$
a	**1**
b	**0**
c	**1**
d	**0**
e	**1**

On the right of each domain element is its image.

End of example

We mention that a relation can be displayed in a similar way, provided that it

doesn't involve a very large number of ordered pairs.

Much of our forthcoming study of Boolean algebra is based on the binary set $\mathbb{B} = \{0, 1\}$. This motivates the following definition.

A *Boolean* function is a function from $\mathbb{B}^n$ to $\mathbb{B}$, where n is some positive integer.

Example:
The following table displays a Boolean function.

(x, y)	$f(x, y)$
$(0, 0)$	1
$(0, 1)$	0
$(1, 0)$	1
$(1, 1)$	0

This function maps $\mathbb{B}^2$ onto $\mathbb{B}$, and the table tells us that $(0, 0)$ has image 1, $(0, 1)$ has image 0, and so on.

To be consistent with our practice of writing the image of a domain element x as $f(x)$, we really ought to write the image of (x, y) as $f((x, y))$. But the usual convention is to drop the outer parentheses, and just write $f(0, 0) = 1$, et cetera.

A more common way to represent a Boolean function like that described in the

table above is to have a separate column for each variable. This gives a table like that shown below.

x	y	$f(x,y)$
0	**0**	**1**
0	**1**	**0**
1	**0**	**1**
1	**1**	**0**

Similarly, a Boolean function from $\mathbb{B}^3$ into $\mathbb{B}$ would have columns headed x, y, z and $f(x,y,z)$; and so on.

End of example

Formative Assessment

You should now do as many practice exercises as necessary to establish that you can correctly establish properties of finite sets.

End of formative assessment

Infinite Sets

We recall the following facts. If X and Y are two finite sets, then:

- $|X| \leq |Y|$ if and only if there is a one-to-one mapping from X into Y;

- $|\boldsymbol{X}| = |\boldsymbol{Y}|$ if and only if there is a one-to-one correspondence betweeen $\boldsymbol{X}$ and $\boldsymbol{Y}$; and
- $|\boldsymbol{X}| < |\boldsymbol{Y}|$ if and only if there is a one-to-one mapping from $\boldsymbol{X}$ into $\boldsymbol{Y}$ but no one-to-one correspondence between them.

Using these ideas we can extend the relations $<$, $\leq$ and $=$ to infinite cardinalities.

Any finite set of cardinality $\boldsymbol{n}$ has the same cardinality as the subset

$$\boldsymbol{\{1, 2, \ldots, n\}}$$

of $\mathbb{N}$. Now since there is clearly a one-to-one map of

$$\boldsymbol{\{1, 2, \ldots, n\}}$$

into $\mathbb{N}$ but no such map onto $\mathbb{N}$, we can say that

$$\boldsymbol{|\{1, 2, \ldots, n\}| < |\mathbb{N}|.}$$

The symbol $\boldsymbol{\omega}$ (lower-case Greek *omega*) is often used to represent the cardinality of the set $\mathbb{N}$. So $\boldsymbol{\omega} = |\mathbb{N}|$. Thus we have

$$\boldsymbol{n < \omega}$$

for every finite number $\boldsymbol{n}$.

It can be shown that $\boldsymbol{\omega}$ is the *smallest* infinite number. (We omit the details.) It is often called "countable infinity", and the set $\mathbb{N}$ is said to be *countably infinite* or *countable*.

Furthermore, any (infinite) set $\boldsymbol{X}$ which can be placed in one-to-one correspondence with $\mathbb{N}$ can be regarded as having the same cardinality as $\mathbb{N}$, and is therefore also countable.

This means effectively that there is a way of writing down the first few elements of $\boldsymbol{X}$ in some kind of order, which establishes an unambiguous pattern for the order of the remaining infinite number of elements (so that there is a way to always know what the "next" element is going to be).

Worked Example:
We show that the set $\mathbb{Z}$ of all integers is countable.

The members of $\mathbb{Z}$ can be listed in the order

$$\mathbf{0, 1, -1, 2, -2, 3, -3, \ldots .}$$

This order defines a mapping

$$\mathbf{0} \text{ to } \mathbf{1, 1} \text{ to } \mathbf{2, -1} \text{ to } \mathbf{3, \ldots}$$

from $\mathbb{Z}$ onto $\mathbb{N}$, where any positive integer $\boldsymbol{n}$ is mapped to the even number $\boldsymbol{2n}$ and any negative integer $\boldsymbol{-n}$ is mapped to the odd number $\boldsymbol{2n + 1}$. This mapping is a one-to-one correspondence between $\mathbb{Z}$ and $\mathbb{N}$, which shows that the two sets have the same cardinality.

End of worked example

This example illustrates a peculiar feature about the relationships between infinite sets. The set $\mathbb{N}$ is a proper subset of $\mathbb{Z}$, and yet they have the same cardinality. In a sense, $\mathbb{Z}$ is no bigger than $\mathbb{N}$, even though $\mathbb{N}$ lies inside $\mathbb{Z}$ and $\mathbb{Z}$ includes elements that are not in $\mathbb{N}$.

It is harder (but not terribly difficult) to show that the set $\mathbb{Q}$ of all rational numbers is countable. However, the correspondence between $\mathbb{Q}$ and $\mathbb{N}$ is not obvious. The rational numbers can be listed in a certain order, but it is not based on their size. Rather it is based on the relationship between their numerators and their denominators when the numbers are expressed as fractions in reduced form. We omit the details.

These results show that *there are just as many positive integers as there are integers, and there are just as many integers as there are rational numbers.*

Exercise:
Show that the set $\boldsymbol{E}$ of all even positive integers is countable.

End of exercise

Does this mean that all infinite sets are equally as big as each other? No! It can be proved that some sets are too big to be

placed in one-to-one correspondence with $\mathbb{N}$. The best-known example is the set $\mathbb{R}$ of all real numbers.

Theorem:
The set $\mathbb{R}$ of all real numbers is uncountable.

End of theorem

Proof:
This is based on Cantor's Diagonal Argument. Assume that there is a list $r_1, r_2, \ldots$ of all real numbers. We construct a real number r as follows. Let it have 0 as its integer part. Make its first fractional digit different from that of r_1, its second different from that of r_2, its third different from that of r_3, and so on. Then by construction this number r must be different to every real number in the list. So r is not **in** the list! This contradicts the assumption that the list was a list of **all** real numbers. So the assumption is false, and it is impossible to create a list of all real numbers.

End of proof

So $\mathbb{R}$ is the first example we've met of an infinite set which is uncountable. The set $\mathbb{R}$ is often called the *continuum*, and its cardinality is called the *continuum number*, denoted by $\boldsymbol{c}$.

So

$$|\mathbb{R}| = \boldsymbol{c}$$

and since there is a one-to-one map of $\mathbb{N}$ into $\mathbb{R}$ but no such one-to-one correspondence, we have that

$$\boldsymbol{\omega < c.}$$

The *continuum hypothesis* asserts that $\boldsymbol{c}$ is the next infinite number after $\boldsymbol{\omega}$. It has been proved to be "undecidable". Loosely speaking, this means that it will always be impossible to know whether or not it is true.

Theorem:
There are countably infinitely many different infinite numbers.

End of theorem

Proof:
This result is based on the observation that a power set $\mathcal{P}(\boldsymbol{X})$ always has larger cardinality than the set $\boldsymbol{X}$ itself does. So there is no largest infinite number. The proof that the number of infinite numbers is countable rather than uncountable is more technical, and is omitted.

End of proof

The infinite number $\boldsymbol{\omega}$ is also denoted by $\boldsymbol{\aleph_0}$, where $\aleph$ is the Hebrew letter *aleph*. So we say that *omega equals aleph nought.*

Because there are countably many infinite numbers, they are often listed as

$$\aleph_0, \aleph_1, \aleph_2, \ldots$$

where $\omega = \aleph_0$. So the continuum hypothesis asserts that the continuum number c is the next infinite number $\aleph_1$. It is certainly true that $c \geq \aleph_1$ but whether c is equal to or strictly greater than $\aleph_1$ is undecidable.

Formative Assessment

You should now do as many practice exercises as necessary to establish that you can correctly determine whether an infinite set is countable or uncountable.

End of formative assessment